A new theory for high definition virtual acoustic display named ADVISE*

Shouichi Takane†, Yōiti Suzuki‡, Tohru Miyajima§ and Toshio Sone†

†Faculty of Systems Science and Technology, Akita Prefectural University, 84–4 Ebinokuchi, Tsuchiya, Honjo, 015–0055 Japan
‡Research Institute of Electrical Communication and Graduate School of Information Science, Tohoku University, 2–1–1 Katahira, Aoba-ku, Sendai, 980–8577 Japan
§Institute of Technology, Shimizu Corporation, 3–4–17 Etchujima, Koto-ku, Tokyo, 135-8530 Japan

(Received 9 December 2002, Accepted for publication 22 May 2003)

Abstract: A new theory for high definition Virtual Acoustic Display (VAD) based on a model called the “Virtual Sphere (VS) model” is introduced in this paper. This method is named ADVISE (Acoustic Display based on the VIrtual SpherE model). In ADVISE, a sphere-shaped boundary is defined around a listener, and the sound transmission from the sound source in the original field to the entrances of the listener’s ears is divided into two parts. One consists of Head-Related Transfer Functions (HRTFs) corresponding to the points on the boundary of the virtual sphere, and the other consists of Room Transfer Functions (RTFs) from the sound source to the points on the boundary. Then, these two kinds of transfer functions are convolved in real-time, taking into account the dynamic changes in these functions due to the listener’s head movement. The Kirchhoff-Helmholtz boundary integral equation is the theoretical basis of this idea. This equation represents that a sound field generated by sound sources outside a certain closed boundary can be synthesized by phantom monopole and dipole sources distributed on the boundary. In this paper, the theory of ADVISE is stated, then the features of ADVISE are described.

Keywords: Acoustic display, Virtual sphere model, Head-Related Transfer Function, Room Transfer Function, Listener’s movement

PACS number: 43.60.Dh, 43.60.Lq, 43.60.Sx, 43.66.Yw [DOI: 10.1250/ast.24.276]

1. INTRODUCTION

Virtual acoustic display (VAD) is one of the strong and promising means to realize auditory display. Numerous studies on the methods for VAD exist [1,2], and those methods are classified into four approaches.

The first approach is so called the multi-channel stereophonic system, the main aim of which is to simulate certain parts of subjective impression of the original sound field, such as presence and spaciousness. This approach, however, seems to be theoretically weak and inexact, and therefore, not useful for high definition auditory display.

In the second approach, the direct and/or reflected sound waves existing in a room are simulated [3]. A practical way to realize this approach is to distribute multiple loudspeakers around a listener. Each loudspeaker reproduces direct and reflected sounds from the corresponding directions. Reverberation is usually reproduced by exponentially decaying impulse responses uncorrelated among the loudspeakers. This type of auditory display is usually combined with the CAD system in architectural acoustics [4].

In the third approach, the theory of sound field reproduction based on the Huygen’s principle [5] or Kirchhoff-Helmholtz boundary integral equation is applied. Some researchers [6–10] reported that the theoretically precise reproduction of sound field can be achieved with this type of VAD. However, a disadvantageous feature of the second and the third approaches is their large-scale apparatus such as vast numbers of loudspeakers needed to implement the sound reproduction through these approaches. This feature detracts portability and thus makes practical implementation difficult.

The fourth approach is based on the binaural or transaural reproduction of sound waves reaching the
entrees of the listener’s external ears. This type of VAD reproduces a sound field including a set of sound images located in specific directions. This can be formed by a relatively small number of loudspeakers, or by headphones [1,2,11–15]. Hence, a portable system can easily be realized not only in laboratory but also in commercial applications [2]. However, the practical application of this type of VAD also involves some problems as follows:

1. Difficulty in synthesis of Room Transfer Functions (RTFs) which involve the direct sound and multiple reflections,
2. Difficulty in dealing with dynamic cues in sound localization, such as listeners’ head movements [15],
3. Variation in Head-Related Transfer Functions (HRTFs) caused mainly by individuality.

As stated above, each type of conventional VAD has both advantages and disadvantages. To overcome the disadvantages of the conventional methods, a concept named the “Virtual Sphere (VS) model” was proposed by the authors [16–18]. In this paper, our proposed theory of the virtual acoustic display based on the VS model is extended to the reproduction in more practical acoustic environments. The theory proposed here for high definition VAD, named ADVISE (Acoustic Display based on the Virtual SpherE model), can realize precise synthesis of HRTFs and Room Transfer Functions (RTFs) theoretically using a practical scale of hardware.

2. THEORY OF ADVISE

2.1. Sound Field Reproduction Based on the Kirchhoff-Helmholtz Boundary Integral Equation

Figure 1 conceptually illustrates the sound field to be controlled. A closed region Ω is established with its boundary Γ. When no sound source exists in Ω, sound pressure at the point \( r_p \) in Ω can be expressed by internal Kirchhoff-Helmholtz boundary integral equation as follows:

\[
P(r_p, \omega) = \int_{\Gamma(r_p)} \left\{ G_F(r_p, r_q, \omega) \frac{\partial P(r_q, \omega)}{\partial n_q} - P(r_q, \omega) \frac{\partial G_F(r_p, r_q, \omega)}{\partial n_q} \right\} d\Gamma_q,
\]

where \( G_F \) is the Green function of the Helmholtz equation. This indicates that the original sound field can be reproduced in Ω if the sound pressure and its derivative on the boundary Γ are equalized with those in the original field. If the boundary Γ is divided into \( N \) small elements \( \Gamma_j (j = 1, \cdots, N) \), and sound pressure as well as its derivative are assumed not to change on these elements, the boundary integral equation can be modified as the equation:

\[
P(r_p, \omega) \approx \sum_{j=1}^{N} \left\{ \frac{\partial P(r_j, \omega)}{\partial n_j} \int_{\Gamma_j(r_j)} G_F(r_p, r_j, \omega) d\Gamma_q - P(r_j, \omega) \int_{\Gamma_j(r_j)} \frac{\partial G_F(r_p, r_j, \omega)}{\partial n_j} d\Gamma_q \right\}.
\]

In the right hand side of Eq. (2), the first term inside the summation can be regarded as the sound pressure from a monopole source distributed on \( \Gamma_j \) whose strength is equal to \( \partial P(r_j, \omega) / \partial n_j \), and the second term indicates the sound pressure from a dipole source located on \( \Gamma_j \) whose strength is equal to \( P(r_j, \omega) \).

It is well known that the Kirchhoff-Helmholtz boundary integral equation is a variation of the Huygen’s principle [19]. The principle for sound field reproduction based on the Kirchhoff-Helmholtz boundary integral equation was proposed by Ise [9], and its improvement by Takane et al. [10] followed.

2.2. Virtual Sphere (VS) Model

The abovementioned principle enables theoretically precise reproduction of the sound field in any enclosed space Ω, under the condition that there is no source inside Ω. The method of ADVISE is introduced based on this principle.

As shown in Figs. 2 and 3, a boundary Γ is set around the listener. While shape of this boundary can be arbitrarily set, a spherical shape is chosen here since it could be convenient for the signal processings relating to the transfer functions (see Sect. 2.4 for detail). This boundary is called a Virtual Sphere (VS) boundary hereafter. Sound transmission system from the sound sources to the listener’s ears is divided into two parts. One is the...
2.3. Synthesis of Transfer Functions

When there exist the \( M \) sound sources outside the virtual sphere boundary, sound pressure at the point \( r_i \) can be expressed as follows:

\[
P(r_i, \omega) = \sum_{j=1}^{M} S_j(\omega) R_j(r_i, \omega),
\]

(3)

where \( S_j(\omega) \) and \( R_j(r_i, \omega) \) indicate the input signal to the \( j \)-th sound source, and RTF (Room Transfer Function) from the point of the \( j \)-th sound source to the point \( r_i \), respectively. The center of the listener’s head is assumed to be the origin of the coordinates. Sound pressure at the listener’s position \( r \) (corresponding to that of listener’s head) inside the VS boundary is obtained from the Kirchhoff-Helmholtz boundary integral equation as the following approximation:

\[
P(r, \omega) \approx \sum_{i=1}^{N} \left\{ A_i(r_i, \omega) \frac{\partial P(r_i, \omega)}{\partial n_i} - B_i(r_i, \omega) P(r_i, \omega) \right\},
\]

(4)

where \( \partial / \partial n_i \) indicates the partial derivative in the direction outward from \( \Gamma_i \), \( A_i(r_i, \omega) \) and \( B_i(r_i, \omega) \) are expressed as

\[
A_i(r_i, \omega) = \int_{\Gamma_i(r_i)} G_F(r, r_i, \omega) d\Gamma',
\]

(5)

\[
B_i(r_i, \omega) = \int_{\Gamma_i(r_i)} \frac{\partial G_F(r, r_i, \omega)}{\partial n_i} d\Gamma,
\]

(6)

where \( G_F(r, r_i, \omega) \) is the Green function of the Helmholtz equation in free field. As stated in Sect. 2.1, Eq. (4) indicates that the sound field inside \( \Omega \) is reproduced precisely by the phantom monopole and dipole sources distributed on \( \Gamma \). Hence, the sound pressure at the listener’s ears can be reproduced by convolving the listener’s HRTFs with each of the phantom sources.

For convenience of implementation, \( A_i(r_i, \omega) \) and \( B_i(r_i, \omega) \) are approximated as follows:

\[
A_i(r_i, \omega) \approx G_F(r, r_i, \omega) \Delta S_i,
\]

(7)

\[
B_i(r_i, \omega) \approx \frac{1}{\Delta_i} \left\{ G_F(r, r_i^+, \omega) - G_F(r, r_i^-, \omega) \right\} \Delta S_i.
\]

(8)

In Eqs. (7) and (8), the integral operation is approximated to the values of the integrand at the center of each element times the area of the \( i \)-th element \( \Delta S_i \), and the derivative operation \( \partial G_F / \partial n_i \) in Eq. (6) is approximated to the difference of \( G_F \) between two closely located points at \( r_i, r_i^+ \) and \( r_i^- \), of which interval is \( \Delta_i \). This corresponds to the approximation of the dipole source to the two closely located monopoles. The sound pressure at the listener’s left and right ears, \( P_L^{\left(0\right)}(\omega) \) and \( P_R^{\left(0\right)}(\omega) \), can be synthesized by using HRTFs for the sound source at \( r_i, r_i^+ \) and \( r_i^- \), when the listener exists at the center of the VS. Here the HRTF corresponds to the “free field transfer function” defined by Blauert [1], which is the transfer function from the sound pressure at a certain measurement point in the auditory canal of the listener to the sound pressure at a point corresponding to the center of the listener’s head in
the absence of the listener. HRTFs for each phantom sound source distributed on the VS are required in ADVISE. From Eq. (4) and those HRTFs $P_L^{(O)}(\omega)$ is expressed as follows:

$$P_L^{(O)}(\omega) = \sum_{i=1}^{N} C_i \left[ \Delta P(r_i, \omega) H_L(r_i, \omega) G_F(r, r_i, \omega) \right.$$

$$\left. - P(r_i, \omega) \left[ H_L(r_i^+, \omega) G_F(r, r_i^+, \omega) \right. \right.$$  

$$\left. - H_L(r_i^-, \omega) G_F(r, r_i^-, \omega) \right].$$

where $H_L$ indicates the left-ear HRTF for the sound source on the VS boundary, $\Delta P(r_i, \omega) = P(r_i^+, \omega) - P(r_i^-, \omega)$, and $C_i = \Delta S_i / \Delta t_i$, $P_R^{(O)}(\omega)$ is obtained from replacing $H_L$ by $H_R$ in Eq. (9). As HRTFs for each phantom sound source on the VS boundary, some databases are publically available [24,25] for example.

2.4. Tracking of the Listener’s Movement

From Eq. (9), it is shown that the change in $P_L(\omega)$ due to the listener’s movement can be traced by changing only the HRTFs for the phantom sources distributed on the VS boundary. If the listener’s movement is restricted to rotation including nodding and pivoting, the distance from the virtual sphere to the listener is regarded as constant. In this case, the HRTFs corresponding to the phantom sources at the virtual sphere are expressed as a function of angles and this markedly simplifies the digital signal processing to trace the real-time change of the HRTFs caused by the head rotation. Thus, the synthesis of the whole sound transmission paths from each sound source to the listener’s ears is extremely simplified. It should also be emphasized as an advantageous feature in ADVISE that the sound pressure in a certain sound field can be simultaneously presented to different listeners with different movement if the set of HRTFs for each listener may be provided. This is because the sound transmission paths from the sound sources to the listeners are expressed by cascading RTFs and HRTFs in ADVISE.

Moreover, listener may move around in the VS if the HRTFs for the phantom sources is obtained. Listener’s position and head rotation can be traced by using magnetic and/or ultrasonic sensors [2]. Since the VS boundary is used only for dividing the sound transmission system into RTFs and HRTFs, size of the VS boundary may be set arbitrarily. However, there exist some limitations. This is discussed in the other paper by the authors [26].

2.5. Discussion on the Application of ADVISE to the Practical Environment

Considering the practical implementation of the VAD based on the abovementioned theory, two sorts of problems arise. One is the synthesis of sound when the sound sources are inside the VS boundary $\Gamma$. Typical example of this case is to synthesize the listener’s own voice. This is impossible within the abovementioned theory, because all sound sources must be outside $\Gamma$. In order to discuss this problem, sound pressure at the listener’s left ear is divided into two components as follows:

$$P_L(\omega) = P_L^{(O)}(\omega) + P_L^{(I)}(\omega),$$

where $P_L^{(O)}(\omega)$ and $P_L^{(I)}(\omega)$ represent the sound pressure generated by the sound sources outside and inside the VS boundary, respectively. Replacing the letter L by R in Eq. (10) yields the expression for sound pressure at the listener’s right ear. $P_L^{(O)}(\omega)$ and $P_R^{(O)}(\omega)$ can be synthesized by the application of the theory stated above.

The other problem is the synthesis of sound from the moving sound sources. The synthesis of sound from the moving sources according to the present theory requires real-time alteration of both the RTFs and the HRTFs.

2.5.1. Synthesis of sound from the sources inside the VS boundary

Synthesis of sound pressure at the listener’s ears can be achieved based on the method mentioned in Sect. 2.3, when the sound sources are outside $\Omega$. However, it is impossible to reproduce the sound pressure by the sound sources in $\Omega$. For the sound sources in the latter case, transfer functions from the moving sources to the listener’s ears must be obtained specifically. If this transfer function from the $j$-th source in $\Omega$ is obtained as $F_{L,j}(\omega)$, $P_L^{(I)}(\omega)$ can be synthesized by computing the following equation:

$$P_L^{(I)}(\omega) = \sum_{j=1}^{M_I} F_{L,j}(\omega) S_{j}^{(I)}(\omega),$$

where $S_{j}^{(I)}(\omega)$ indicates the source signal of the $j$-th source, and $M_I$ is the number of the sources in $\Omega$. $P_R^{(I)}(\omega)$ is computed in the same way. If the size of the VS is small, number of such sources may be relatively small in practical implementation, thus the consideration of the sound sources inside the VS may not cause any serious increase in the amount of processings.

2.5.2. Synthesis of sound from the moving sources

The transfer functions from the moving sources to the listener change dynamically by the movement of both the listener and/or the sources. If the RTFs of the moving sources are obtained by RTFs for the moving sound sources inside the VS boundary, $F_{L,j}(\omega)$ and $F_{R,j}(\omega)$ $(j = 1, \cdots, M_I)$ may also change dynamically due to the movement of both the source and the listener. Synthesis of sound from the sources moving across the VS boundary, for example, may be difficult to deal with by the VS model because the
processing of those sources is different, and the processing must be altered depending on their position. When the original sound field includes such sources, transfer functions from those sources to the listener should be synthesized independently of the other sound sources.

2.6. Reproduction System

The theory of ADVISE is utilized to synthesize the sound pressure at the listener’s ears. Hence a system to reproduce the synthesized sound pressure with ADVISE may be arbitrary such as the binaural and the transaural systems [20,21]. The system must have, however, the real-time traceability of the dynamic change in HRTFs due to the listener’s movement. As for the transaural systems, it should be noted that when the crosstalk cancellation is executed, the dynamic change in the transfer functions from the loudspeakers to the listener’s ears caused by the movement of the listener in the reproduction environment should also be traced in real-time [20,22].

3. OVERVIEW OF ADVISE

3.1. Formulation using the vector and matrix notations

To express the overall algorithm of ADVISE, the processings introduced in the previous section is described in terms of the vector and matrix notations.

1. Source signals of the static sources and the corresponding RTFs from them to the points on the VS boundary are convolved. Here, the vector consisting of the source signals $S_j(\omega)(j = 1, \ldots, M)$, and the matrix filter of RTFs of which component is represented by $R_j(r, \omega)$ are expressed as $s$ and $R$ respectively. The matrix filters of RTFs which have the components $R_j(r^+ \omega)$ and $R_j(r^- \omega)$ are also needed to compute the derivatives of the sound pressure on the VS boundary. These are expressed by $R^+$ and $R^-$ respectively. When the symbols $P$, $P^+$ and $P^-$ are used for the vectors composed of $P(r, \omega)$, $P(r^+ \omega)$ and $P(r^- \omega)$ respectively, the following equations are obtained:

$$ P = Rs, \quad P^+ = R^+s, \quad P^- = R^-s. \quad (12) $$

When the number of the phantom sources on the VS boundary is $N$, the multiplication and addition of the order $O(\text{MN})$ are required.

2. Eq. (9) is described in terms of the vector expressions as follows:

$$ P_L^{(0)}(\omega) = \Delta P^T(h_L g) - P^T(h_L^+ g^+ - h_L^- g^-), \quad (13) $$

where $\Delta P = P^+ - P^-$, $h_L$, $h_L^+$, $h_L^-$, $g$, $g^+$ and $g^-$ are vectors consisting of $H_L(r, \omega)$, $H_L(r^+ \omega)$, $H_L(r^- \omega)$, $C_i G_F(r, r_i, \omega)$, $C_i G_F(r, r_i^+, \omega)$ and $C_i G_F(r, r_i^-, \omega)$ respectively. Terms $h_L^+ g^+$ and $h_L^- g^-$ in Eq. 13 respectively indicate the product of each component of corresponding two vectors. The similar expression for the sound pressure at the listener’s right ear is given as the following:

$$ P_R^{(0)}(\omega) = \Delta P^T(h_R g) - P^T(h_R^+ g^+ - h_R^- g^-). \quad (14) $$

In Eqs. 13 and 14, the components of the vectors $h_L$, $h_L^+$, $h_L^-$, $h_R$, $h_R^+$, $h_R^-$ are dynamically changed due to the listener’s movement. The multiplication and addition of the order $O(N^2)$ are required.

3. The following vector products yield the sound pressure at the listener’s ears when the static sound sources are situated inside the VS boundary:

$$ P_L^{(l)}(\omega) = f_L^T s^{(l)}, \quad (15) $$

$$ P_R^{(l)}(\omega) = f_R^T s^{(l)}, \quad (16) $$

where $f_L$ and $f_R$ are vectors whose components are $F_L(\omega)$ and $F_R(\omega)$, and $s^{(l)}$ is a vector whose component is $S_j^{(l)}(\omega)$ $(j = 1, \ldots, M_l)$. Vectors $f_L$ and $f_R$ may vary due to the listener’s movement. Operation counts for this part are $O(M_l^2)$.

4. Total sound pressure at the listener’s ears is obtained by summing the components computed in 2. and 3., i.e., Eq. (10).

Although the formulation in this paper was made in the frequency domain, the algorithm listed above can be implemented in the time domain by taking inverse Fourier Transforms of the transfer functions and the source signals, and computing the convolutions of them instead of the products in the frequency domain. When the processings listed above are executed in time domain, the order of the operation counts becomes $O(N^3\text{HRTF}\text{RTF})$, where $L_{\text{HRTF}}$ and $L_{\text{RTF}}$ are respectively the sample length of a HRTF and a RTF. Here the number of sources inside the VS boundary is assumed to be small comparing to $N$. Consequently, the operation counts may be huge comparing to those in the conventional VADs. A primary real-time system developed by the authors, which is stated in another paper [26], has five DSPs of maximum speed of 1 GFLOPS each. Sound pressure at the listener’s ears in the 2-dimensional sound field up to 1 kHz can be synthesized by using that system.

The synthesis of the sound pressure at the listener’s ears in the arbitrary sound field with ADVISE requires the matrices of RTFs: $R$, $R^+$, $R^-$; and the vectors of HRTFs: $h_L$, $h_L^+$, $h_L^-$, $h_R$, $h_R^+$, $h_R^-$. For HRTFs of dummy heads and certain subjects, there exist some databases measured in anechoic rooms [24,25]. Moreover, the estimation of HRTFs using Boundary Element Method (BEM) is discussed [27–30]. As for RTFs, numerical analysis of sound field by geometrical methods [21], Finite Element Method [31] and BEM [32,33] seems practical.
Obtaining RTFs for the VS boundary based on measurement may be hard.

3.2. Block Diagram of ADVISE

According to the processings described in Sect. 3.1, block diagram of ADVISE is illustrated in Fig. 4. In this figure, the source signals are assumed to be divided into two classes according to whether the sources are inside or outside the VS boundary. The VS model is adopted in the processings of the sources outside the VS boundary. This means that $P_L^{(O)}(\omega)$ and $P_R^{(O)}(\omega)$ are synthesized from the sound pressure and its derivative at the points on the VS boundary, no matter how many sources are distributed. Adjustment for the change in the HRTFs corresponding to the points on the VS boundary is needed. Adjustment of the change in RTFs due to the movement of the listener and the sources is also required.

When the number of the sources outside the VS boundary is small as compared with that of the phantom sources distributed on the VS boundary, the processing cost in ADVISE is estimated much higher than those in the conventional auditory displays, in which the transfer functions from those sources to the listener are directly convolved with the source signals. However, it is difficult and not practical to obtain the set of transfer functions in the original sound field for the individual listener. Moreover, the RTFs may change abruptly by a slight movement of the listener’s positions [23]. In contrast, in the VS model of ADVISE the synthesis of the transfer functions is divided into RTFs and HRTFs, and the real-time processing corresponding to the listener’s movement is required only for the synthesis of the HRTFs for the points on the VS boundary.

As for the sound sources inside the VS boundary, change in the transfer functions from the sources to the listener must be adjusted. The number of sources of this kind may be small as compared with the sources outside the VS boundary, thus the processings needed in the synthesis of the sound from those sources may not seriously affect the total processing costs.

3.3. Other Processings in ADVISE

There exist some other processings required in ADVISE which do not explicitly appear in Fig. 4. To adjust the change in the transfer functions due to the movements of the listener and the sources, the transfer functions must be revised in real-time. Since these transfer functions are not obtained as their functional forms, their changes are...
adjusted by the interpolation in the functions measured (or computed) at discrete points. As for the HRTFs for the points on the VS boundary, some databases of HRTFs open in public are available \([24,25]\). However, not only \(h\) but also \(\hat{h}^+\) and \(\hat{h}^-\) are also needed in the VS model. How to get approximations of \(\hat{h}^+\) and \(\hat{h}^-\) from \(h\) is discussed in another paper by the authors \([26]\). The interpolation of the transfer functions must be processed both spatially and temporally, and the same way as that used in the conventional auditory displays \([2]\) is available.

4. CONCLUDING REMARKS

In this paper, a new method for VAD named ADVISE was introduced. It was theoretically shown that accurate reproduction of sound waves reaching the entrances of the listener’s ears was realized with ADVISE.

ACKNOWLEDGMENT

This study was carried out under the Corporate Research Project of Res. Inst. Electr. Comm., Tohoku University (H14/A06), and a part of this study was supported by the Grant-in-Aid for Scientific Research (Subject No. 12650359, 14350193 and 13650419).

REFERENCES

Shouichi Takane was born in Tokyo, Japan in 1967. He received B. E., M. E., and Ph. D. degrees from Tohoku University in 1989, 1991 and 1998, respectively. He is currently an associate professor at the Faculty of Systems Science and Technology, Akita Prefectural University. His research interests include high-definition acoustic display, sound field control, numerical analysis of sound field and acoustic signal processing. He is a member of the Acoustical Society of Japan, the Virtual Reality Society of Japan, the Institute of Electronics, Information and Communication Engineers, and other societies.

Yoichi Suzuki graduated from Tohoku University in 1976 and received his Ph. D. degree in electrical and communication engineering in 1981. He is currently a professor at the Research Institute of Electrical Communication, Tohoku University. His research interests include psychoacoustics, high-definition auditory display and digital signal processing of acoustic signals. He received the Awaya Kiyoshi Award and Sato Prize from the Acoustical Society of Japan. He served as a vice-president of the Acoustical Society of Japan from 2001 to 2003 and is currently the Editor in Chief of the Society.

Tohru Miyajima was born in Fukuoka, Japan in 1963. He received B. E. degree from Kyushu Institute of Design in 1986. He joined Institute of Technology, Shimizu Corporation in 1986. He has been engaged in research on room acoustics, noise control, and evaluation of environmental sound. He is a member of the Acoustical Society of Japan and the Architectural Institute of Japan.

Toshio Sone graduated from Tohoku University in 1958 and received his Ph. D. degree in electrical and communication engineering in 1963. He was a professor of Tohoku University until 1999 and is now a professor of the Faculty of Systems Science and Technology, Akita Prefectural University. He has been engaged in researches on psychological acoustics, electroacoustics, room acoustics, and noise control for more than 40 years. He is a former president of the Acoustical Society of Japan and a fellow of the Acoustical Society of America, a member of the Institute of Noise Control Engineering/Japan, a member of the Engineering Academy of Japan and other societies.